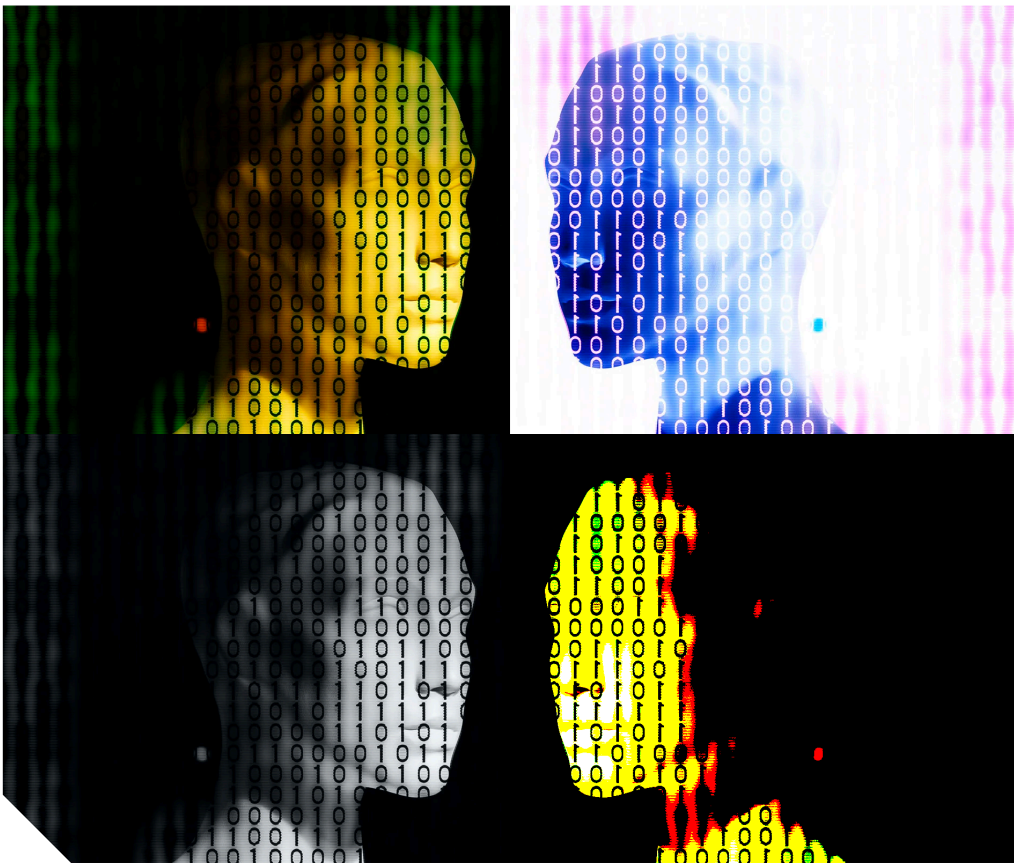

How to Ensure Data Integrity and Maintain It

WHITE PAPER

Web-Hed Technologies, Inc. (dba Webhead)
1710 N. Main Ave San Antonio, Texas 78212 (210) 354-1661 webheadtech.com

Data is centric to the growth and success of every organization. The data generated across different layers of business operations can determine the organization's performance and help decision-makers better understand the current market environment.

Given the value of data to the growth of businesses and organizations, it is critical to have measures that guarantee data integrity. The business and its decision-makers draw insight into the operation, success, viability, and expansion of the organization.

What Is Data Integrity?

Data integrity is defined as the trustworthiness and reliability of data from generation, collection, transfer, storage, backup analysis, and utilization. It also refers to the safety of the data and its compliance with regulatory standards, ensuring its accuracy, completeness, and consistency.

When data integrity is secured, the data collected remains complete, accurate, and reliable regardless of the storage period.

Although data integrity goes hand-in-hand with data security, the two are different. Data security revolves around limiting data access, while integrity is more encompassing.

Why Is Data Integrity Important?

The quality, reliability, and consistency of the data you use to make decisions will determine the success of the organization and the impact of the decision you make. Data integrity ensures searchability and traceability of the data to its source for verification.

Integrity also ensures effective data accuracy and protection. Compromised data is of no use to most companies and leads to ill-advised decisions.

Taking proper data integrity precautions saves the organization from costly data audit trails required to trace errors and recover the data.

With the ever-increasing volume and value for data, data integrity not only protects the data but, by extension, the organization and company. Leaking or compromised data will adversely affect the performance and reputation of the organization.

According to a study by IBM, the average cost of a data breach to a U.S. organization is $8.6 million and it takes up to 280 days to identify and contain the breach. It is far more cost-effective to put measures in place to ensure the integrity of the data than to deal with the aftermath.

Factors Affecting Data Integrity

Several areas can compromise integrity in various stages of the lifecycle of data. A few examples include:

Human error – Erroneously entering data, deleting, duplicating, or failing to follow proper protocols is the leading risk.

Transfer errors – This error occurs when data cannot be effectively or completely transferred from one database location to another. A transfer error can also occur when some data is present in the destination location but not in the source or vice versa.

Bugs and viruses – Spyware, viruses, and malware can infiltrate the computer and erase, change or steal information, compromising the authenticity and accuracy of the data.

Compromised hardware – Sudden server crashes, compromised hardware performance, or sudden hardware failure can damage data, limit its access or remove some of the data, making it difficult to use.

Types of Data Integrity

Ensuring and maintaining data integrity takes many forms depending on the stage of the data lifecycle.

Physical Integrity

This is the protection of the accuracy and wholeness of the data while being stored or retrieved. Physical integrity is compromised when a natural calamity hits and power goes out. Hackers can also interrupt database functionalities core to the storage and retrieving of data.

When the physical integrity of the data is compromised, the data is inaccessible, inaccurate, or has a variety of other difficulties.

Logical Integrity

Logical integrity aims to keep the data intact as it is being used in various ways. It protects the data from human mistakes and hackers and is often divided into four categories:

Entity integrity

As a primary rule of effective database construction, Entity integrity is the process of enforcing database table primary keys where keys must be with unique non-null values assigned to single rows or groups of rows.

Referential integrity

Referential integrity refers to the accuracy between related tables. All database tables are related to one another, so one table's primary key can appear in another table. Referential integrity is each primary key's dependency on foreign keys.

Domain integrity

Domain integrity requires that all relational database columns be declared on a specific domain. Data items are the primary units in the relational data model.

User-defined integrity

In User-defined integrity, users themselves create rules and constraints customized to fit their particular requirements. This form of integrity is employed when other types of integrity are inadequate to safeguard data.

How to Ensure Data Integrity

With the increasing automation and computerization of systems and data collection, the expectations and regulations on data integrity are at an all-time high. Here are some recommendations that will help you maintain integrity within your systems:

Data Entry Training

Human error is among the leading causes of data integrity risks, largely because most employees don't know how to preserve or enter the data. Training employees on how to enter and maintain data is a great first step to ensuring the integrity and accuracy of data. Instilling a sense of responsibility among the employees to preserve data quality will also ensure that everyone is hands-on in preserving and maintaining data integrity.

Select Appropriate System and Service Providers

Data integrity is so sensitive that it is comprehensively covered in the FDA's Title 21 Code of Federal Regulations Part 11. Any system and service provider with access to your data should be fluent in the relevant regulations and must be fit-for-purpose. Before hiring any new companies to handle your data, validate the efficacy of their systems and learn about their organizational culture and maturity in data management.

Validate Input and Data

If your data is being supplied by an unknown source like an end-user or some other application, it is imperative to verify and validate the data inputs to ensure their accuracy. Data validation should also be a regular process executed at intervals to ensure that data processes have not been corrupted.

Remove Duplicate Data

Duplicate data causes ambiguity in the system, causing malicious errors and data integrity breaches. Identifying and removing duplicate data within a reasonable time is critical. Large organizations have teams dedicated to cleaning duplicate files.

Depending on the volume of data or the size of the organization, you can also try this approach or use software to help you remove the duplicate data. However, these are not always accurate, so you might need to validate them after the process.

However, some of the tools used for the process are open source. While they might be cost effective, they can further compromise data integrity. Only use tools that meet regulation standards.

Back Up the Data

While taking measures to preserve data integrity, you cannot afford to lose data permanently. Regularly backing up data is critical in ensuring you can restore everything to normal in case of a cybersecurity attack or hardware failure.

Although cloud storage solutions are the best backup option, take caution when choosing providers to minimize risk and only go for options that provide you with the best security services.

Access Controls

When you have no access control in your organization, you're at a higher data integrity risk. Individuals with malicious intent can easily gain access to your database and cause grievous harm to the organization.

The least privilege model is the best approach. Only reserve access to the core users and ensure a high level of control to preserve data integrity.

Keep an Audit Trail

Whenever you have a data breach, it's crucial to track its source. Keeping an audit trail will preserve data integrity.

Audit trails provide organizations with tidbits that indicate the source of the problem for effective action and resolution.

Good audit trails should have a record of all data in the system, which includes changes made to the database or individual files. A good audit trail should:

- Be tamper proof and cannot be manipulated by the users.
- Generate the audit trails automatically.
- Track and record every event on the database and files.
- Align with the users, so you know who accessed the audit trail.
- Have timestamps for all events.

It is good practice to audit your audit trails regularly to identify any loopholes or areas of improvement to better safeguard data integrity.

Collaboration in the Organization

All members of your organization should be on the same page when it comes to the value and seriousness of data integrity. They should know who is making data changes and when and what is expected from each of them. If there are protocols, they need to know what they are and how they impact the company and the data they handle.

Encrypting Data

Data encryption is an effective measure to preserve data integrity within the organization. Even if someone has access to the data, they cannot read it without the decryption key.

Data encryption is a vital data security measure that protects your data when attackers get access to files stored in the database by stealing the hardware. Webhead can encrypt your data.

Pay Attention to Cybersecurity

Cybersecurity plays a vital role in preserving data integrity in every organization. Observing proper cybersecurity protocols will limit access to essential data without permission, reduce and prevent chances of a data breach.

A comprehensive cybersecurity approach includes having strict policies, using advanced security tools, and taking the necessary measures to guarantee the integrity of your data. Webhead has professionals who can scan your data or website's coding to look for gaps and vulnerabilities, and close the gaps.

The Bottom Line

Collecting data for organizations is the easy part these days. Preserving data integrity is the more challenging part. Having proper strategies, tools, and policies in place are critical in ensuring you have accurate and authentic data to empower you to make the right decision.

With the increase in cybercrime and sometimes honest human mistakes, it is possible for other organizations and individuals to use your data to work for them or mislead you. Contact us to get an expert's view on how you can secure your data integrity and use your data to your advantage.